Status of LIGO Data Analysis System

Albert Lazzarini

LIGO Laboratory at Caltech

LSC Meeting

14 March 2001

Baton Rouge, LA
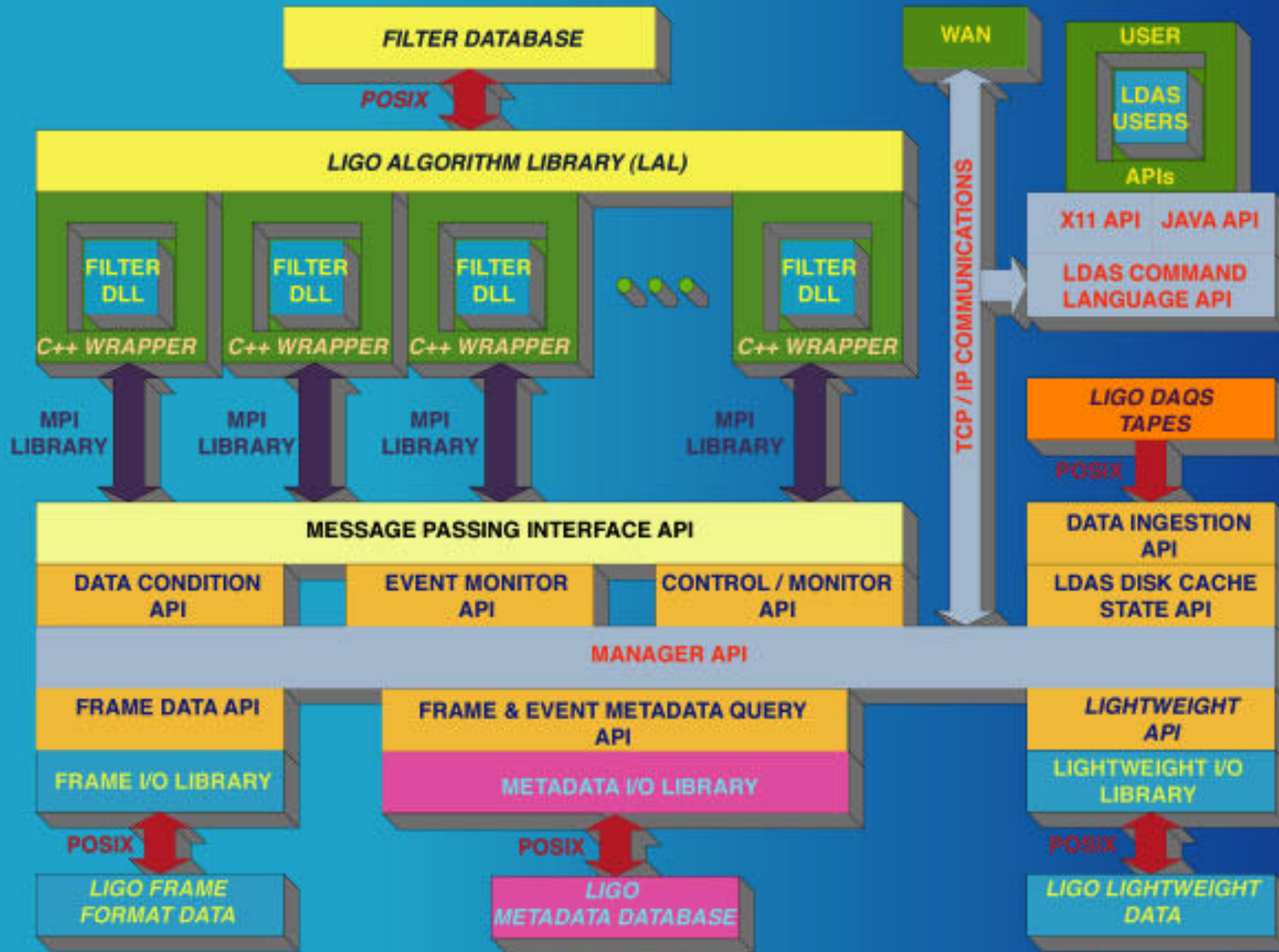
LDAS Update for LSC 2001/03

- **Software:**
  - » LDAS code base is >600,000 lines of C++ and Tcl/Tk now -- will grow to ~1,000,000 lines by time of LIGO I Science Run.
  - » Work on 14 of 16 APIs has commenced
    - – Final two APIs (disk cache, data ingestion) awaiting decision on archive
    - – Will use LDAS interface to APIs available with chosen solution
  - » 1 Year ago: focus on code development
  - » Now: focus on MDCs, engineering run support
    - – Software integration
    - – Complete code development
    - – Add functionality as required through MDC experiences

LDAS Update for LSC 2001/03

- ## Hardware:
  - » LDAS main procurement has begun
  - » Phase I: issued 9 Mar 2001
    - – 28 TB of RAID disk systems for Observatories, Caltech, MIT
      - Sites will have 1 month look-back capacity on spinning media @3 MB/s per interferometer
      - Caltech HPSS disk cache
      - MIT disk cache
    - – 6000 slot (500+TB capacity) robotic silo for HPSS at Caltech
    - – PCs, servers for E6 run
  - » Phase II: -- after final beowulf, HPSS benchmark tests
    - – Data movers for HPSS (support 5 continuous streams of data at full bandwidth)
    - – PCs for main clusters at Caltech, MIT, LLO, LHO (total of ~400 units)

LIGO-G010093-00-E

LDAS Update for LSC 2001/03

- ## Hardware:
  - » Phase III: complete HPSS for Science run
    - – STK model 9940 tape drives for HPSS
    - – Tapes for Science run
    - – Start build-up of large disk farm in front of HPSS for data analysis
      - Continue to grow farm throughout Science run
      - Keep up with data growth
      - Target keeping all commonly used, needed data on disk
      - Use HPSS for backup, large data dumps to disk caches

LIGO-G010093-00-E

Plan to Reach Science Run
Detector & Data Analysis

- **Jan to mid-March**
  - *LHO 2k, continued work on improving robustness of lock, some work on sensitivity*
  - *LLO 4k, Lock single arm, recombined Michelson with Fabry-Perot (F-P) arms, Power Recycled Michelson (PRM)*
  - *LHO 4k, installation*
  - » ***SW: Prepare LDAS release for E3***
  - » ***HW: Procure Phase I, final RAID configurations, HPSS tape silo, small beowulf clusters for E6***

- **February 28: Seattle Earthquake**

- **March 9-12**
  - *E3 (engineering run): coincidence run between LHO PEM and single F-P arm at LLO*
  - » ***SW+HW: Archive E3 data***

Plan to Reach Science Run
Detector & Data Analysis

- mid-March to mid-May
  - *LHO 4k, complete installation, lock mode cleaner*
  - *LHO 2k, repair, suspension sensor replacement, resurrect PRM studies*
  - *LLO 4k, lock full interferometer, sensitivity/robustness*
    - » ***MDC: metaDataAPI (Caltech/LHO)***
    - » ***SW: Prepare LDAS release for E4***
    - » ***HW: Install Phase I hardware***
    - » ***HW: Benchmark HPSS on Sun hardware at Sun testbed facilities, Beaverton, OR***

- May
  - *E4 run: LLO 4 km only, operating in recombined mode (possibly recycling)*
    - » ***MDC: MPI inspiral search (first of 4 MDCs tied to upper limits run)***
    - » ***SW+HW: Archive E4 data***

LIGO-G010093-00-E

Plan to Reach Science Run

Detector & Data Analysis

- **May - June**
  - *LHO 2k, bring full interferometer back on-line, sensitivity studies*
  - *LLO 4k, improve full interferometer lock, sensitivity studies*
  - *LHO 4k, PRM locking (no arms yet)*
  - » ***SW: Prepare LDAS release for E5***
  - » ***HW: Specify HPSS HW configuration for Phase II of procurement***

- **late June - early July**
  - *E5 LHO 2k and LLO 4k in full recycled configuration,*
    LHO 4k in PRM mode, probably poorer sensitivity due to earthquake delay
  - » ***MDC: Stochastic background search***
  - » ***SW+HW: Archive E5 data***

- **July - Sept**
  - *LLO 4 k suspension sensor replacement, bring back on-line*
  - *LHO 2km sensitivity studies, 4k lock full interferometer*
  - » ***SW; Prepare LDAS release for E6***
  - » ***MDC : Burst search***
  - » ***HW: Benchmark PCs for large beowulf procurement, Procure Phase II***

Plan to Reach Science Run

Detector & Data Analysis

- **late Sept**
  - *E6 triple coincidence run with all 3 interferometers in final optical configuration ("upper limit run")*
  - » ***SW+HW: Archive E6 data, on-site upper limit searches***

- **Oct – early 2002**
  - *Improve sensitivity and reliability*
  - *Alternate diagnostic testing with engineering runs*
  - » ***MDC : Data Archive***
  - » ***HW: Specify HPSS drives, tapes, IDE/SCSI RADI 5 disk cache for data at Caltech; Procure Phase III***

- **Jan - July 2002**
  - » ***SW+HW: Prepare Release 1 of LDAS for Science run***
  - » ***SW+HW : Integrated single interferometer running on-site***
  - » ***SW+HW : multiple interferometer running off-site***

LIGO/LSC Software Deployment Schedule
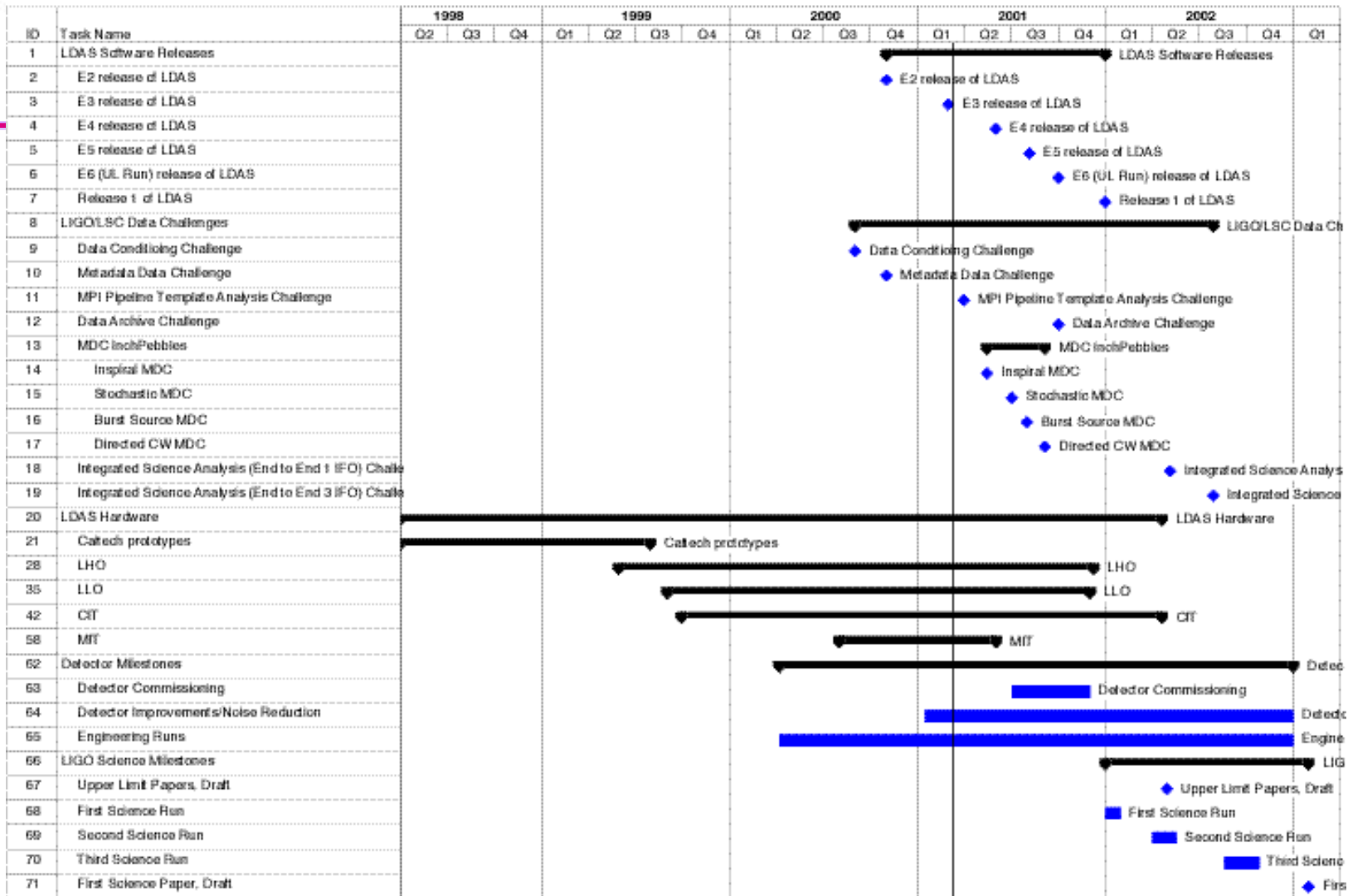
LDAS: The next generation

- **Data Analysis White Paper revision**
  - » Anderson, Lazzarini, Nash, Riles, Weinstein, Weiss, Wiseman (chair),

- **LIGO:**
  - » 2Q: Participating with CACR in Distributed Teraflop Facility proposal
  - » 3Q: MRE proposal to NSF for Adv. LIGO; need to identify/cost LDAS upgrades (LIGO Laboratory) for Advanced LIGO interferometers

- **LSC Resources**
  - » GriPhyN -- ITR2000 -- SW/CS component
    - – Caltech/UWMUTB participation
  - » International Virtual Data Grid Laboratory (iVDGL) -- ITR2001
    - – Proposal due 25 April 2001
    - – Funds for multiple Tier2 centers for GriPhyN research
    - – EU involvement in distributed computing (Italy/UK)

GriPhyN
Grid Physics Network

- **Caltech**
  - » Working on virtual data models with USC/Information Science Institute (Carl Kesselman/ISI)
    - – Consider how to develop new LDAS APIs for accessing LIGO data from the grid environment
      - • Staging large computational tasks
      - • Tracking data for delivery over a distributed grid environment

- **UWM**
  - » New CS/physicst hired jointly with NCSA to work on importing grid tools to the UWM beowulf
  - » Use of beowulf systems to mirror large datasets ("datawulf")

GriPhyN
Grid Physics Network

- **UTB**
  - » Hired scientist to head the educational outreach program of GriPhyN

    - Tutorials,
    - Workshops
    - Web pages
    - Etc.